

TESTING A NEW ANALYTIC MODEL FOR GRAVITATIONAL LENSING PROBABILITIES

SUDEEP DAS AND JEREMIAH P. OSTRIKER
 Princeton University Observatory, Princeton, NJ 08544-1001
Draft version February 5, 2008

ABSTRACT

We study gravitational lensing with a multiple lens plane approach, proposing a simple analytical model for the probability distribution function (PDF) of the dark matter convergence, κ , for the different lens planes in a given cosmology as a function of redshift and smoothing angle, θ . The model is fixed solely by the variance of κ , which in turn is fixed by the amplitude of the power spectrum, σ_8 . We test the PDF against a high resolution Tree-Particle-Mesh simulation and find that it is far superior to the Gaussian or the lognormal, especially for small values of $\theta \ll 1'$ and at large values of κ relevant to strong lensing. With this model, we predict the probabilities of strong lensing by a single plane or by multiple planes. We find that for $\theta \sim 10''$, a single plane accounts for almost all ($\sim 98\%$) of the strong lensing cases for source redshift unity. However, for a more typical source redshift of 4, about 12% of the strong lensing cases will result from the contribution of a secondary clump of matter along the line of sight, introducing a systematic error in the determination of the surface density of clusters, typically overestimating it by about 2 – 5%. We also find that matter inhomogeneities introduce a dispersion in the value of the angular diameter distance about its cosmological mean. The probable error relative to the mean increases with redshift to a value of about 8% for $z \simeq 6$ and $\theta \sim 10''$.

Subject headings: cosmology: theory — gravitational lensing — large-scale structure of the universe — methods: analytical

1. INTRODUCTION

Gravitational lensing provides a unique tool for studying the cosmological distribution of matter, because it directly probes the gravitational potential and is indifferent to the nature or the physical state of matter. It is, therefore, free from the assumptions that can plague other dynamical methods for studying gravitational fields.

Tidal deflections of light rays by structures along its path can weakly distort and magnify distant sources. These effects, although small on the individual level, can be statistically significant for an ensemble of sources. Thus, statistical properties of the intervening matter distribution can be inferred from such *weak lensing* studies (see, for e.g. Blandford & Narayan 1992; Bartelmann & Schneider 2001; Refregier 2003, for reviews). This field is extremely promising at the present time but is fraught with observational challenges, because these subtle effects have to be very delicately isolated from comparable systematic errors.

Occasionally, the distortion due to the intervening matter may be strong enough to produce multiple images of the source. *Strong lensing* is extremely sensitive to the cosmology and hence, can be used to constrain cosmological parameters. Small changes in the cosmological parameters can significantly alter the nonlinear structure formation history of the universe and thereby result in huge differences in the abundance of compact objects which behave as potential strong lenses. This makes the probability of any specific type of strong lensing event a useful observable for constraining cosmology (see, for example, Turner et al. 1984; Narayan & Blandford 1991; Kaiser 1992; Cen et al. 1994; Nakamura & Suto 1997; Cohn et al. 2001; Keeton 2001; Premadi et al. 2001; Oguri

et al. 2002; Li & Ostriker 2002, 2003). For example, the statistics of lensed quasars or galaxies lensed by clusters into giant arcs can be used to study the population of lenses, their distribution in redshift space, mass, and hence the cosmological model. Strong lensing systems may also be studied individually, in order to obtain a detailed understanding of the structure and the distribution of matter in the lensing object and the source (see, for example, Claeskens et al. 2000; Broadhurst et al. 2000; Shapiro & Iliev 2000; Cohn et al. 2001; Belokurov et al. 2001; Patnaik & Narasimha 2001; Inada et al. 2003; Fassnacht et al. 2005; Sluse et al. 2005; Wayth et al. 2005; Claeskens et al. 2006). In this paper, we shall be mostly interested in the former problem, namely the statistics of strong lensing events.

In the absence of satisfactory analytic tools, the most widely used method for studying strong lensing statistics has been the analysis of large numerical simulations. The method of tracing the path of light rays through a simulated universe evolving in accordance with some specific cosmological model is being extensively used to study the multitude of lensing phenomena. For example, giant arc statistics using ray tracing simulations of clusters in a Λ CDM cosmology are being thoroughly investigated (examples include Bartelmann et al. 1998; Wambsganss et al. 2004a; Dalal et al. 2004; Ho & White 2004; Li et al. 2005; Hennawi et al. 2005a,b, etc.). However, these simulations are costly in that they require extensive computing power and time. This also restricts the range of cosmological models that can be investigated; in most cases only the currently favored cosmological model is used. One must not, however overlook the array of important analytical works from the so-called halo model approach (see, for e.g. Li & Ostriker 2003; Oguri et al. 2003; Oguri & Keeton 2004), which allow accurate com-

putation of lensing probabilities for a single lens.

The goal of this paper is to provide an analytic method to address some of the questions pertaining to gravitational lensing statistics that, at the cost of having much more limited information available than from detailed numerical studies, would have the benefit of allowing any definite cosmological model to be investigated. We test our results against numerical simulations performed by the Tree-Particle-Mesh (TPM) code (Bode & Ostriker 2003) for a Λ CDM universe. The method lends itself easily to the treatment of lensing under other cosmologies.

We consider the fluctuations in the matter density along a light cone of a given opening angle. We propose an analytic method to compute the probability that a light beam will encounter a given surface density while passing through a slice of matter centered around a given redshift. The model is completely defined by the variance of the projected density of the matter in the slice, smoothed on the angular scale corresponding to the beam. We show how the variance of surface mass density on the lens can be predicted from the power spectrum in Section 2 and develop the analytic model in Section 3.

Armed with this result, we adopt a multiple lens plane approach to study strong lensing. We divide up the line of sight distance to the source into contiguous slices of equal comoving depth, each of which is chosen to be larger than the correlation length so that each slice can be considered statistically independent of the others. We imagine the matter inside each of these slices to be projected onto a plane at its center. The probability density function (PDF) for the surface density on each plane is then constructed with our analytical method. The matter encountered by the light ray while propagating through the universe is then simulated by randomly picking a surface mass density from the PDF on each slice it passes through. In fact, the surface density scaled by the so called critical density required for multiple image formation, better known as the convergence, is computed on each of these planes. We identify different strong lensing scenarios according as whether the individual or the additive value of convergences so obtained is in excess of unity. We shoot a statistically significant number of rays to obtain the probabilities of different events. In particular, we study the probability of strong lensing as a function of source redshift and angular scale.

Another important question is the role of auxiliary planes aiding the main lens to produce multiple images. It is known that some multiply imaged quasars (e.g. the quadrupole system B1422+231, see Kormann et al. 1994) cannot be well modeled under the thin lens approximation, which assumes all mass responsible for lensing to be concentrated on a mathematical surface of zero thickness (Schneider et al. 1992). Also, observations suggest cases where galaxies at disparate redshifts appear to have contributed to the lensing event (Tonry 1998; Augusto et al. 2001; Chae et al. 2001; Johnston et al. 2003; Fassnacht et al. 2005). In order to study the probabilities of such events, we identified, within the strong lensing cases, the ones that were caused by a single plane or by the concordance of two or more planes and studied their probabilities separately. Incidentally, a similar study was done in Wambsganss et al. (2004b), who studied the phenomenon at the pixel level ($\sim 1.5''$) and reported a signif-

icant contribution of auxiliary planes in the strong lensing probabilities. With our method, we can conveniently include an element of angular smoothing depending on the strong lensing situation and study the probabilities as a function of smoothing radius. We treat this problem in Section 4.

We also note that our model works very well even in the weak lensing regime (angles $\gtrsim 1'$). The statistics of weak lensing convergence can be studied from the magnification effects of clustered matter which leads to variations in the number density and image sizes of galaxies across the sky (see, e.g., Jain 2002). Also, shear maps generated from weak lensing surveys yield statistical properties of the convergence. Thus a good analytic model for the convergence can be useful in constraining the properties of large scale structure (Barber et al. 2004), although the practicalities involved may seem rather daunting at the moment (Bartelmann & Schneider 2001). The situation will improve with forthcoming surveys having high signal-to-noise ratios, enabling us to study the small scale clustering of matter.

In Section 5, we study the effect of the chance accumulation of secondary matter along the line-of-sight on the estimation of cluster masses and finally, in Section 6, we examine the dispersion in the angular diameter distances due to inhomogeneous distribution of matter.

2. VARIANCE IN PROJECTED OVERDENSITIES

Since the fluctuations in matter density along the line of sight parallel the growth of structure in the universe, with higher redshifts corresponding to an earlier era in the structure formation history, these variances are intricately related, and can, in fact, be predicted from the evolution of the matter power-spectrum. For large scales or early times the growth of structure is linear and the power-spectrum is linearly related to its primordial form through a transfer function (see Bardeen et al. 1986). However, on small scales and late times, the growth of structure becomes nonlinear and small scale modes couple and grow more rapidly than in the linear regime. The evolution of the power-spectrum in this era cannot be calculated analytically with sufficient accuracy and one has to resort to numerical simulations to find some analytic fitting function for the simulated power spectrum. This was first done by Peacock & Dodds (1996) who used N-body simulations based on hierarchical clustering of matter to produce an analytic fit. Later, Smith et al. (2003) proposed a halo model (Seljak 2000; Ma & Fry 2000) inspired fitting function which was shown to be in excellent agreement with numerical simulations (Barber & Taylor 2003) in the context of the lensing convergence power-spectrum. We have adopted the Smith et al. (2003) fitting functions in this work. In order to calculate the variances in the 2D matter distribution obtained by projecting the matter along the line of sight in a redshift slice, the 3D power-spectrum has to be related to its 2D counterpart. This is done by the well known Limber approximation (Kaiser 1998). Using this method, we show that the 2D variances and consequently, the probability distribution for surface density can be predicted with sufficient accuracy given a cosmological model.

In the present approach, we divide the matter along the line of sight from the observing redshift of zero to a redshift of 6.34 into 38 slices, each $160 \text{ h}^{-1}\text{Mpc}$ thick,

and project the matter in each of these slices onto a plane at its center. Therefore we have lens planes at $80 + j \times 160 \text{ h}^{-1}\text{Mpc}$ for $j = 0, \dots, 37$. Let $\Sigma(\xi, z)$ denote the projected surface mass density at a position $\xi = \{\xi_1, \xi_2\}$ on a plane at redshift z . We denote the same quantity, when smoothed by a Gaussian window of angular radius θ_0 , by $\Sigma_{\theta_0}(\xi, z)$. We are interested in finding the variance $\sigma_2^2(\theta_0, z) \equiv \langle \delta_2^2(\xi; \theta_0) \rangle$ in the projected overdensities $\delta_2(\xi)$ given by,

$$\delta_2(\xi; \theta_0) = \frac{\Sigma_{\theta_0}(\xi, z) - \langle \Sigma_{\theta_0} \rangle}{\langle \Sigma_{\theta_0} \rangle}. \quad (1)$$

Here and henceforth, the subscript 2 on any symbol is meant to assert that it is a quantity depending on 2-dimensional variables. In order to obtain $\sigma_2^2(\theta_0, z)$, we first use the Limber Equation (Kaiser 1998) to obtain the 2-D power spectrum for the distribution of Σ on a plane from the nonlinear 3D power spectrum $P_{\text{NL}}(\mathbf{k})$. Let us denote the comoving distance to the center of the bounding planes of the i th slice by χ_{i-1} and χ_i , so that the first slice is bounded planes at by χ_0 and χ_1 and so on. The coordinate of a point inside the i th slice can be denoted by,

$$\mathbf{x}(\chi, \boldsymbol{\theta}) = \chi(z)(\theta_1, \theta_2, 1), \quad (\chi_{i-1} < \chi < \chi_i),$$

where θ_1 and θ_2 are two dimensionless numbers (angles) specifying the position of the point on the plane orthogonal to the line of sight. Since we are interested only in the projection of the matter inside this slice, the overdensity seen at the point (θ_1, θ_2) on the plane of projection is,

$$\delta_2(\boldsymbol{\theta}) = \int d\chi W(\chi) \delta[\mathbf{x}(\chi, \boldsymbol{\theta})], \quad (2)$$

where W is a selection function, which, in our case, is given by a step function,

$$W(\chi) = \begin{cases} \frac{1}{(\chi_i - \chi_{i-1})} & \text{if } (\chi_{i-1} < \chi < \chi_i) \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

The Fourier space conjugate to $\boldsymbol{\theta}$ is denoted by \mathbf{l} , and defined through,

$$\delta_2(\mathbf{l}) = \int d^2\boldsymbol{\theta} e^{i\mathbf{l}\cdot\boldsymbol{\theta}} \delta_2(\boldsymbol{\theta}). \quad (4)$$

The 2D power-spectrum is given by the standard Limber approximation (Kaiser 1998; Dodelson 2003),

$$P_2(l) = \int d\chi \frac{W^2(\chi)}{\chi^2} P_{\text{NL}}(l/\chi, z) \quad (5)$$

$$= \frac{1}{(\chi_i - \chi_{i-1})^2} \int_{\chi=\chi_{i-1}}^{\chi_i} d\chi \frac{P_{\text{NL}}(l/\chi)}{\chi^2}, \quad (6)$$

using our selection function.

By introducing $k \equiv l/\chi$, we can rewrite this as,

$$P_2(l, z) = \frac{1}{l(\chi_i - \chi_{i-1})^2} \int_{l/\chi_i}^{l/\chi_{i-1}} dk P_{\text{NL}}(k, z). \quad (7)$$

Now, the variance of the projected overdensity, Gaussian smoothed at an angle θ_0 can be calculated by using,

$$\sigma_2^2(\theta_0, z_i) = \frac{1}{(2\pi)^2} \int \Lambda^2(l\theta_0) P_2(l) d^2\mathbf{l}, \quad (8)$$

where $\Lambda(l, \theta_0)$ is the Fourier transform of the $\boldsymbol{\theta}$ space Gaussian smoothing window of width θ_0 , and is given by,

$$\Lambda(l, \theta_0) = e^{-l^2\theta_0^2/2}.$$

With this we have,

$$\sigma_2^2(\theta_0, z_i) = \frac{1}{(2\pi)^2} \int d^2\mathbf{l} e^{-l^2\theta_0^2} P_2(l) \quad (9)$$

$$= \frac{1}{2\pi(\chi_i - \chi_{i-1})^2} \int dl e^{-l^2\theta_0^2} \int_{l/\chi_i}^{l/\chi_{i-1}} dk P_{\text{NL}}(k, z_i) \\ = \frac{\pi}{(\chi_i - \chi_{i-1})^2} \int dl e^{-l^2\theta_0^2} \int_{l/\chi_i}^{l/\chi_{i-1}} \frac{dk}{k^3} \Delta_{\text{NL}}^2, \quad (10)$$

where $\Delta_{\text{NL}}^2 = \frac{1}{2\pi^2} P_{\text{NL}}(k) k^3$ is the dimensionless power spectrum.

The integral above requires the value the non-linear power spectrum $\Delta_{\text{NL}}^2(k)$ as a function of scale k and redshift z . Incidentally, this can be generated using the routine “halofit” by Smith and Peacock (see Smith et al. 2003).

We test our predictions against a high resolution cosmological simulation using the Tree-Particle-Mesh (TPM) code (Bode & Ostriker 2003). The simulations were performed in a box with a comoving side length of $L = 320h^{-1}\text{Mpc}$. We used $N = 1024^3 = 1,073,741,824$ particles, so the individual particle mass is $m_p = 2.54 \times 10^9 h^{-1} \text{ M}_\odot$. The cubic spline softening length was set to $\epsilon = 3.2h^{-1} \text{ kpc}$, producing a ratio of box size to softening length of $L/\epsilon = 10^5$. The output was stored at 19 redshift values out to $z \approx 6.4$, such that the centers of the saved boxes matched comoving distances of $(160+n \times 320)h^{-1}\text{Mpc}$, where $n = 0, \dots, 18$. Each box was then divided into 9×9 separate square cylinders along its length. This was done for three orthogonal projections, leading to 243 sub-volumes. Two lens planes were produced for each sub-volume by bisecting along the line-of-sight and projecting the mass in each $160h^{-1}\text{Mpc}$ -long volume onto a plane. Thus, we have planes at comoving distances of $(80 + j \times 160)h^{-1}\text{Mpc}$, where $j = 0, \dots, 37$. At the highest redshift, each plane has a side length of $35.6h^{-1}\text{Mpc}$ and contains 800^2 pixels, making the pixel size $44.4h^{-1}\text{kpc}$ comoving. At lower redshifts, an opening angle of about $20''$ was maintained by keeping the number of pixels constant but progressively decreasing the size of the planes. Thus in the lowest redshift box the lensing planes are $1.9h^{-1}\text{Mpc}$ on a side and the pixel size is $2.3h^{-1}\text{kpc}$, which is slightly lower than our spatial resolution. Because at lower redshifts the plane does not cover all of the corresponding sub-volume, there is a random offset perpendicular to the line of sight within each sub-volume. In all lensing applications discussed below, we will not consider structures smaller in size than 5 pixels or three times our spatial resolution.

In order to test the results against the WMAP normalized (Spergel et al. 2003) ΛCDM TPM simulations, we used the same values of cosmological parameters as used by the TPM code, to generate the power spectrum by halofit. These parameters are, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_b = 0.04$, dimensionless Hubble constant, $h = 0.7$, $\sigma_8 = 0.95$, and the derived power-spectrum shape parameter (Sugiyama 1995) $\Gamma \equiv \Omega_m h \exp \left[-\Omega_b \left(1 + \frac{\sqrt{2h}}{\Omega_m} \right) \right]$.

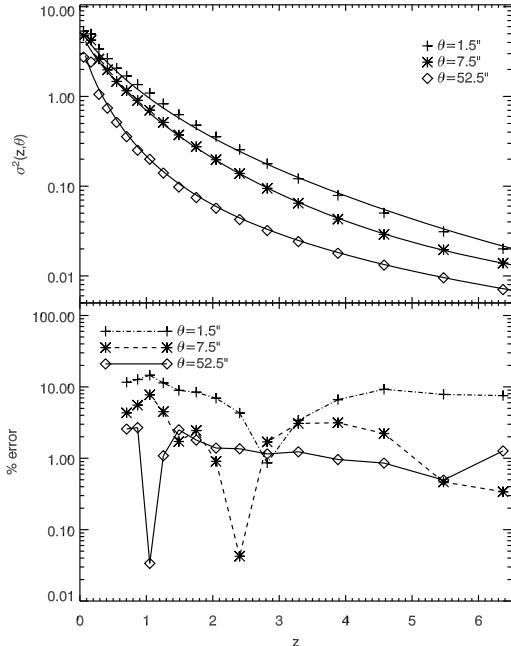


FIG. 1.— *Top:* Comparison of the analytic result for the variance $\sigma_2^2(z, \theta)$ as a function of redshift z for three different smoothing angles θ , with the data obtained from the quoted TPM numerical simulations (Bode & Ostriker 2003). The points are the variances from the N-body simulations. The lines represent the analytic result obtained by using the Limber equation and the nonlinear power spectrum of Smith et al. (2003). *Bottom:* The percentage errors in the predicted values of the variance

We used redshifts corresponding to the center of each of the $160 \text{ h}^{-1}\text{Mpc}$ thick slices to generate the power spectrum for that slice. Then we used the Limber Approximation and the Gaussian filtering by the desired angle [equation (10)], to obtain the variances of the 2-D overdensities on the planes at each of these redshifts. The results are displayed in Fig. 1. This test shows that we are able to predict the variance within each of these slices from theory to roughly 1–10% accuracy.

3. THE MODEL

The probability distribution function for the lensing convergence for large angular scales, $> 10'$, can be straightforwardly obtained from analytic perturbative calculations (Stebbins 1996; Villumsen 1996; Bernardeau et al. 1997; Jain & Seljak 1997; Kaiser 1998; Schneider et al. 1998; van Waerbeke et al. 1999). However, on angular scales less than $5'$, the perturbative approach fails and non-linear gravitational clustering comes into play. Hence, a number of models have been developed to move from a linear to the fully nonlinear power spectrum, and from there to the PDF of the convergence field. On intermediate scales ($1 - 10'$) such analytical models have been shown to have good predictive power when compared with numerical simulations (see Barber et al. 2004, and references therein.). Recently, Kayo et al. (2001) analyzed the one- and two-point statistics of the three dimensional dark matter distribution using *N*-body simulations and found them to be well described by the lognormal model (Kayo et al. 2001). Following this, Taruya et al. (2002) have shown that the lognormal model, with certain parameterizations, can fairly

describe the convergence field for small angular scales $2' \lesssim \theta \lesssim 4'$ and up to a level threshold $\nu \sim 10$, where $\nu \equiv \kappa / \langle \kappa^2 \rangle^{1/2}$. However, satisfactory analytic methods to generate the convergence PDF on the ultra non-linear scales, $< 1'$, are as yet unknown. Here, we put forward a phenomenological model for the PDF of the surface mass density and hence, that of the convergence on the lens sheets, which seems to work down to few arcseconds and is designed to better approximate the high density tail of the distribution, which reduces to a distribution akin to the lognormal and then to the Gaussian for large angles and high redshifts. The model is completely defined by the variance of the surface mass density on each sheet and thus encodes all higher order statistical information in the two point statistics of Σ .

We computed the probability density function (PDF) of the quantity $x = \frac{\Sigma}{\langle \Sigma \rangle} \equiv 1 + \delta_2$ on each plane. It is found that the PDF is well described by the following form, inspired by the lognormal distribution,

$$f(x) = \frac{N}{x} \exp \left[-\frac{(\ln x + \omega^2/2)^2 (1 + A/x)}{2\omega^2} \right], \quad (11)$$

where the values of the three parameters, A , ω and N are fixed by the following three constraints,

$$\text{Normalization : } \int_0^\infty f(x) dx = 1, \quad (12)$$

$$\text{By defn. } \langle x \rangle = 1 : \int_0^\infty x f(x) dx = 1, \quad (13)$$

$$\text{Observed variance: } \int_0^\infty (x - \langle x \rangle)^2 f(x) dx = \sigma_2^2. \quad (14)$$

Therefore the model is fixed by a single input, the variance $\sigma_2^2(\theta, z)$ of the projected overdensity of at a given redshift and a given angular scale of smoothing. Since this quantity can be theoretically predicted, as discussed in the previous section, the model is essentially fixed by the cosmology.

A search algorithm is developed which, given a value of σ_2^2 , returns the values of the three parameters satisfying all above constraints.

The result is displayed in Fig. 2, where the parameters A , ω^2 and N are plotted as functions of $\sigma_2^2(x)$. Note that all these parameters are approximately power-laws of σ_2^2 for small values of the latter.

The theoretical model is shown against the PDF obtained from simulations for angular smoothing radii of $7.5''$, $52.5''$ and $1.5'$ in Figs. 3, 4 and 5, respectively, with the lognormal model plotted over each. A detailed comparison of the model and other standard PDFs are displayed in Fig. 6.

We would like to stress here that although our model has three variables, A , ω and N , it is not a three parameter fit in the sense that the three variables are not varied to obtain some kind of “best” fit with the data points. We have simply put forward a non-standard probability distribution which is on the same footing as the Gaussian or the lognormal as far as the number of free parameters is concerned. All the three distributions have their variables determined by the fact that they are normalized, have unit mean by construction and have a certain variance predicted by the cosmology. Thus, each of the three distributions is essentially fixed by the input cosmology and the knowledge of the redshift and the smoothing angle. In Figs. 3, 4 and 5 we have simply overlaid the theoretical model on the data points; there is no fitting with the data involved anywhere.

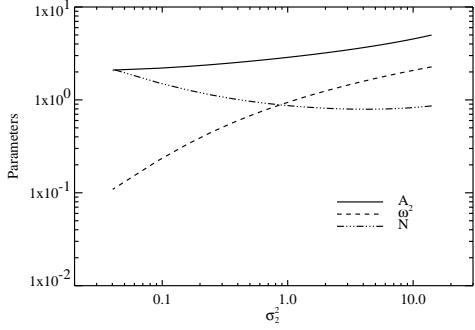


FIG. 2.— The parameters A , ω^2 and N plotted as a function of $\sigma_2^2(x)$ (cf. equation (11))

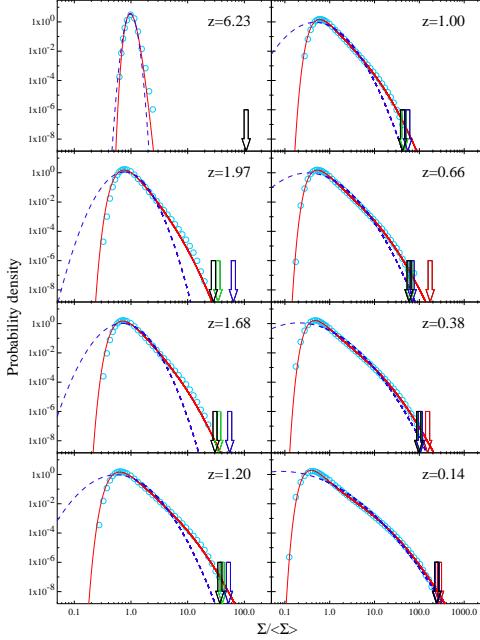


FIG. 3.— The theoretical model (red line) shown against the probability densities (blue circles) obtained from the quoted TPM simulation (Bode & Ostriker 2003) for different redshifts. The dashed dark blue line is the lognormal model. The smoothing angle for this set of data was $7.5''$. The color coded arrows show the position of the critical surface density Σ_{crit} in units of $\langle \Sigma \rangle$, for four different source redshifts: 7.0 (black), 5.0 (green), 3.0 (blue) and 1.0 (red).

It is immediately apparent from these figures that our model does better than the lognormal or the Gaussian in describing the PDF of the surface density, in particular it better describes the all important tail of the distribution which characterizes the low probability but observationally important high κ events. The lognormal hugely underestimates the probability density for just those rare events in the tail of the distribution which are important for strong lensing. This effect is more pronounced for smaller smoothing angles, as evident from Fig. 3, where we show, by arrows, the critical surface densities required to produce strong lensing, for four representative source redshifts. At these values, the probability densities predicted by the lognormal model are orders of magnitude lower than those obtained from the TPM simulation. For all smoothing scales, it appears that on the right tail of

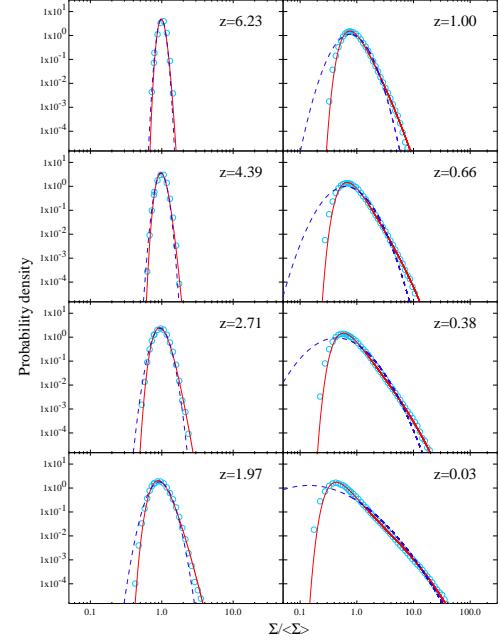


FIG. 4.— The theoretical model (red line) shown against the probability densities (blue circles) obtained from the quoted TPM simulation (Bode & Ostriker 2003) for different redshifts. The dashed dark blue line is the lognormal model. The smoothing angle for this set of data was $52.5''$.

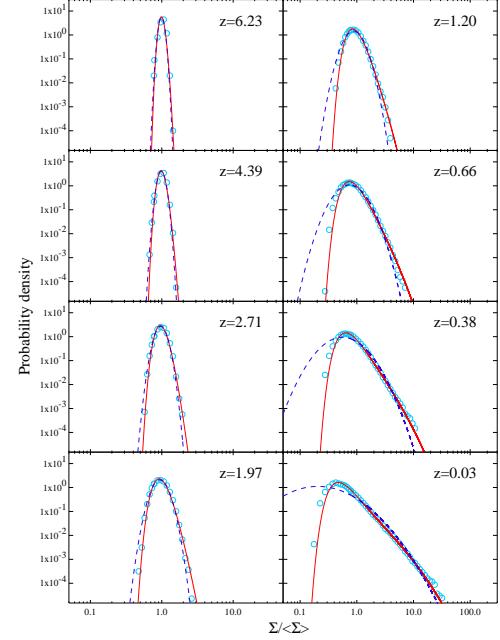


FIG. 5.— The theoretical model (red line) shown against the probability densities (blue circles) obtained from the quoted TPM simulation (Bode & Ostriker 2003) for different redshifts. The dashed dark blue line is the lognormal model. The smoothing angle for this set of data was $90''$ or $1.5'$.

the distribution, the lognormal does moderately well as long as probability densities are above $\sim 10^{-2}$, below which it departs from the data points. Our model, however, appears to describe these heavy tails acceptably. Thus the lognormal proves to be a poor approximation for the very high surface density tails, except for very low

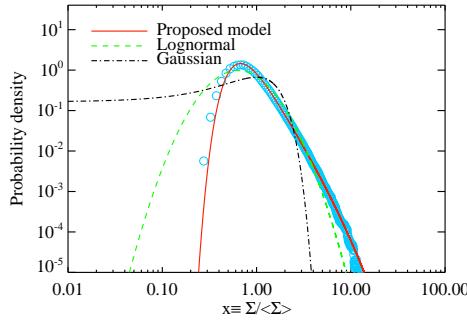


FIG. 6.— Comparison of the lognormal and the Gaussian probability density functions with the model proposed in the paper. The data used for this illustration is for a redshift of 0.7 and a smoothing angle of 52.5''.

or very high redshifts. In the latter limit, both of these models tend to the Gaussian, as will be shown below, and become almost indistinguishable. Another aspect of the distribution that the lognormal fails to describe completely is its skewness. The simulated distributions fall sharply toward the low surface density end, whereas the lognormal hugely overestimates the probability density in this region. The proposed model appears to capture this feature of the distribution with appreciable accuracy. It goes without saying that the Gaussian fails in all respects, except at high redshifts and of course, for large enough smoothing angles for which the underlying Gaussian initial spectrum of perturbations remains dominant. (cf. Fig. 6).

Now, we shall discuss some limiting behavior of the above model. For compactness, we shall drop the subscript 2 in σ_2 in the following discussion. Let us start by considering the exponent in the above formula, namely,

$$\frac{(\ln x + \omega^2/2)^2(1 + A/x)}{2\omega^2} \quad (15)$$

Since, ω^2 and A appear to be well behaved functions of σ^2 , one can write,

$$\begin{aligned} \omega^2 &= c_0\sigma^2 + O(\sigma^4) \\ A &= a_0 + a_1\sigma^2 + O(\sigma^4) \end{aligned} \quad (16)$$

We can expand each component of equation (15) as,

$$(\ln x + \omega^2/2)^2 = (\ln x)^2 + c_0(\ln x)\sigma^2 + O(\sigma^4) \quad (17)$$

$$1 + \frac{A}{x} = 1 + \frac{a_0}{x} + \frac{a_1}{x}\sigma^2 + O(\sigma^4) \quad (18)$$

Therefore, the argument of the exponential can be expanded as,

$$\begin{aligned} \frac{(\ln x + \omega^2/2)^2(1 + A/x)}{2\omega^2} &= \frac{1}{2c_0\sigma^2} \left[(\ln x)^2 \left(1 + \frac{a_0}{x}\right) \right. \\ &\quad + \left\{ a_1 \frac{(\ln x)^2}{x} + c_0 \left(1 + \frac{a_0}{x}\right) \ln x \right\} \sigma^2 \\ &\quad \left. + O(\sigma^4) \right] \end{aligned} \quad (19)$$

For $\sigma^2 \ll 1$ up to order $O(\sigma^{-2})$, this reduces to,

$$\frac{(\ln x + \omega^2/2)^2(1 + A/x)}{2\omega^2} = \frac{1}{2c_0\sigma^2} (\ln x)^2 \left(1 + \frac{a_0}{x}\right) \quad (20)$$

For small σ^2 , the range of x over which the exponent remains appreciable, is also small,

$$x = 1 + \delta$$

where $\delta^2 \sim O(\sigma^2)$. In this approximation,

$$\ln x = \ln(1 + \delta) \simeq \delta$$

and

$$\frac{1}{x} = 1 - \delta.$$

This implies,

$$\frac{(\ln x + \omega^2/2)^2(1 + A/x)}{2\omega^2} \simeq \frac{\delta^2}{2c_0\sigma^2} (1 + a_0) \quad (21)$$

Under these approximations the model reduces to,

$$\begin{aligned} \frac{N}{x} \exp \left[-\frac{(\ln x + \omega^2/2)^2(1 + A/x)}{2\omega^2} \right] &\simeq N \exp \left[-\frac{\delta^2}{2c_0\sigma^2} (1 + a_0) \right] \\ &= N \exp \left[-\frac{\delta^2}{2\beta^2\sigma^2} \right] \end{aligned} \quad (22)$$

Where we have defined $\beta \equiv (1 + a_0)/c_0$. One can easily verify that the condition in equation (14) gives, $\sigma^2 = \sigma^2\alpha^2$, implying $\alpha = 1$. So the normalization condition, equation (12) gives us $N = 1/(\sqrt{2\pi}\sigma)$. This is also graphically illustrated in Fig. 7. Thus our model correctly reduces to a Gaussian for small values of variance.

For small σ , we have $1 + a_0 = c_0 \simeq 3$ and therefore,

$$\omega^2 \simeq 3\sigma^2 + O(\sigma^4)$$

At this point, it is important to note that the surface

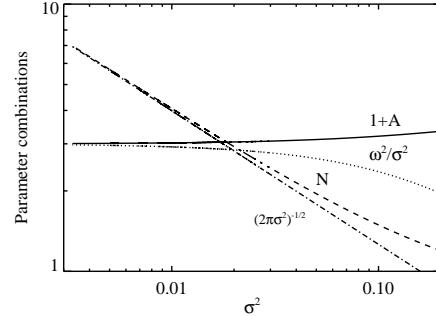


FIG. 7.— At small σ^2 , the parameters $(1 + A)$ and $\frac{\omega^2}{\sigma^2}$ become identical implying $(1 + a_0) = c_0 \simeq 3$ (cf. equation (16)). Also, the normalization parameter N tends to the Gaussian value $1/\sqrt{2\pi}\sigma$. These imply that the distribution tends to the Gaussian form with variance σ^2 for small values of the latter.

density, Σ , that we have been considering so far, contains in it the contribution from the smooth background matter density $\langle \Sigma \rangle$. However, lensing is sensitive only to the departures in the surface mass density from the average rather than its total value (Petters et al. 2001, equations 3.32 and 3.48). One should note in this connection that the values of Σ relevant to strong lensing events are usually much higher than the background density $\langle \Sigma \rangle$. This often leads to neglecting the subtraction of background density in the definition of κ when discussing strong lensing. In many cases, this is indeed a very small effect. However, we have found appreciable differences,

especially in the multiple plane lensing probabilities (see Section 4), if this correction is not taken into account. Therefore, we define the lensing surface mass density Σ_l , that usually appears in the lensing equations as,

$$\Sigma_l \equiv \Sigma - \langle \Sigma \rangle. \quad (23)$$

For the purpose of studying the possibility of strong lensing by a plane, it is important to notice that a multiple image of the source results if the beam encounters a background-subtracted surface density, Σ_l along its path which is in excess of the critical surface mass density $\Sigma_{\text{crit}}(z)$ at that redshift, which is given by,

$$\Sigma_{\text{crit}}(z) = \frac{c^2}{4\pi G} \frac{D_s}{D_l D_{ls}}, \quad (24)$$

where D_s , D_l and D_{ls} are the angular diameter distances from the observer to the source, the observer to the lens, and the lens to the source, respectively.

In order to address the problem in which one, two or more planes together become supercritical, it is convenient to change the variable of interest from $x = \frac{\Sigma}{\langle \Sigma \rangle}$ to the convergence,

$$\kappa = \frac{\Sigma_l}{\Sigma_{\text{crit}}} = \frac{\langle \Sigma \rangle}{\Sigma_{\text{crit}}} (x - 1). \quad (25)$$

The PDF $g(\kappa)$ for the convergence is related to that of x , $f(x)$ by a simple Jacobian transformation,

$$g(\kappa) = \alpha f(\alpha\kappa + 1), \quad (26)$$

with, $\alpha = \frac{\Sigma_{\text{crit}}}{\langle \Sigma \rangle}$.

A physical meaning to the variable α can be attached by defining the minimum value κ_{min} of the convergence κ , corresponding to an “empty” beam, with $\Sigma = 0$. In this case, equation (23) implies $\Sigma_l = -\langle \Sigma \rangle$ and hence,

$$\kappa_{\text{min}} = -\frac{\langle \Sigma \rangle}{\Sigma_{\text{crit}}} = -\frac{1}{\alpha}. \quad (27)$$

Therefore, one may write $\alpha = \frac{1}{|\kappa_{\text{min}}|}$ and the PDF for the convergence can be rewritten as,

$$g(\kappa) = \frac{1}{|\kappa_{\text{min}}|} f\left(\frac{\kappa}{|\kappa_{\text{min}}|} + 1\right), \quad (28)$$

or, written out in its entirety as,

$$\begin{aligned} g(\kappa) d\kappa &= N \\ &\times \exp\left[-\frac{(\ln(1+\kappa/|\kappa_{\text{min}}|) + \omega^2/2)^2 \{1 + A/(1+\kappa/|\kappa_{\text{min}}|)\}}{2\omega^2}\right] \\ &\times \frac{d\kappa}{\kappa + |\kappa_{\text{min}}|}. \end{aligned} \quad (29)$$

The constraint equations, (12), (13) and (14), translate to,

$$\text{Normalization : } \int_{\kappa_{\text{min}}}^{\infty} g(\kappa) d\kappa = 1, \quad (30)$$

$$\text{By defn. } \langle \kappa \rangle = 0 : \int_{\kappa_{\text{min}}}^{\infty} \kappa g(\kappa) d\kappa = 0, \quad (31)$$

$$\text{Observed variance: } \int_{\kappa_{\text{min}}}^{\infty} \kappa^2 g(\kappa) d\kappa = \langle \kappa^2 \rangle. \quad (32)$$

This transformation leads us to the possibility of expanding the applicability of our model beyond the thin mass-sheet case that we have been considering so far. Although we will not consider this here, it may be interesting to ask how much improvement this model provides over the

lognormal form studied by Taruya et al. (2002) in describing the line-of-sight or effective convergence, given by (see, e.g. Bartelmann & Schneider 2001),

$$\kappa_e = \frac{3}{2} \Omega_m \left(\frac{H_0}{c}\right)^2 \int_0^{\chi_s} d\chi g(\chi, \chi_s) \delta(\chi), \quad (33)$$

where χ is the comoving distance, χ_s is the comoving distance to the source, $\delta = \Delta\rho/\langle \rho \rangle$ is the density contrast, and

$$g(\chi, \chi_s) = r(\chi) \frac{r(\chi_s - \chi)}{r(\chi_s)} (1 + z), \quad (34)$$

$r(\chi)$ being the comoving angular diameter distance. This will be the subject of a future paper.

4. STRONG LENSING PROBABILITIES

The necessary condition for a strong lensing situation is $\kappa_e \simeq \sum \kappa_i > 1$, where i is the index of a plane in front of the source. As the convergence κ at a point on each plane can be either positive or negative, a supercritical plane does not necessarily imply a supercritical line of sight. In fact, as Premadi & Martel (2004) note, underdense regions tend to occur around cluster lenses. However, when a single plane or a number of planes, collectively, is supercritical, the probability that negative convergences on some of the remaining planes will make the line of sight subcritical is actually small. Neglecting this real but small correction for the moment, we shall show that there exists a possible analytical calculation to compute the probability that a single plane produces strong lensing. Then, in a following subsection we shall turn to a more accurate numerical (Monte Carlo) approach.

Notation: Let \mathcal{F}_i denote the probability that the i -th plane is supercritical, i.e.

$$\mathcal{F}_i \equiv \text{Prob} [\Sigma(z_i) > \Sigma_{\text{crit}}(z_i)] = \int_1^{\infty} g(\kappa(z_i)) d\kappa(z_i). \quad (35)$$

4.1. Probability that one and only one plane is supercritical

This is equal to the probability that the first plane is supercritical and all the other planes are not, or the second plane is supercritical and the other planes are not, and so on:

$$\begin{aligned} \mathcal{P}_{\text{one}} &= \sum_{i=1}^N \mathcal{F}_i \prod_{\substack{j=1 \\ j \neq i}}^N (1 - \mathcal{F}_j) \\ &\simeq \sum_{i=1}^N \mathcal{F}_i (1 - \sum_{j \neq i} \mathcal{F}_j), \end{aligned} \quad (36)$$

where N is the number of planes in front of the source, and the last step follows from the fact that $\mathcal{F}_i \ll 1, \forall i$. An extension of this scheme for two or more planes becomes increasingly cumbersome and impractical as the number of planes increases. Also, as we already pointed out, this method does not correct for the fact that some lines of sight may become subcritical due to negative convergences on the planes other than those considered in the probability calculation. Therefore, we turn to the more efficient Monte Carlo approach to study these probabilities.

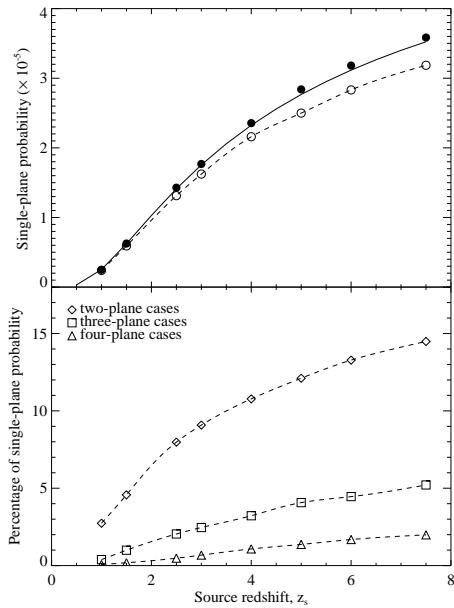


FIG. 8.— *Top:* The probability, as a function of source redshift, that a single plane alone renders the line of sight supercritical. The smoothing angle for this figure was $7.5''$. The solid line is the analytical approximation (cf. equation (36)). The filled circles represent the single-plane probabilities generated by a Monte Carlo simulation which does not take into account the correction discussed in the text. The open circles represent the same, but with the correction applied. As discussed in the text, the analytical calculation slightly overestimates the true probabilities (open circles) as it includes rare cases for which, although the single plane is supercritical, the line of sight is rendered subcritical by negative convergence on some other plane or planes. The open circles are joined by a spline (dashed line) as a guide to the eye. *Bottom:* The probabilities, shown as the percentage of the corresponding single-plane values, that two, three, or four planes together render the line of sight supercritical. The points are from the Monte Carlo study and are joined by splines (dashed lines) to aide the eye.

4.2. Monte Carlo method.

We randomly draw values, κ_i , from the PDFs of κ on the planes lying in front of the source and consider only those cases for which the total convergence is greater than unity. With $\sum \kappa_i > 1$ satisfied, if the highest value among the κ_i 's is greater than unity we identify this as a single plane strong lensing case. If this fails, we consider the two highest values of κ from our crop and look at the sum of the two. If the sum exceeds unity then we identify this as a two plane lensing case. We extend this method to include three or more planes.

4.3. Results

As we have already shown, the probabilities for the single-plane cases are easy to compute analytically (cf. equation (36)), as long as we do not care whether underdensities on the other planes are making the line of sight subcritical. With this assumption, the analytical calculation slightly overestimates the probabilities as expected. However, it is useful as a quick and handy way to get an upper limit to the probabilities. The results from the Monte Carlo study do take the effect of underdense planes into account. The results from both the analytical and the Monte Carlo approach, for a smooth-

ing angle of $7.5''$, are displayed in Fig 8. We note from the figure that when the constraint $\sum \kappa > 1$ is not applied in the Monte Carlo study and cases with $\kappa_{\max} > 1$ but $\sum \kappa < 1$ are wrongly counted as legitimate single-plane cases, the single-plane probabilities obtained analytically match those from the Monte Carlo simulation, as expected. When the correction is taken into account, the points from the Monte Carlo study lie slightly below the analytical values, because we lose a few cases due to the fact that the line of sight was subcritical although the single plane was supercritical. We notice that for source redshifts, $z_s \simeq 1.0$, the contribution of secondary clumps of matter to strong lensing is almost negligible ($\sim 2-3\%$ of the single-plane cases). However, for larger redshifts secondary planes become more and more important and contribute up to about 15% of the single-plane cases for $z_s = 6.0$. Our results are somewhat lower than those obtained in a previous study by Wambsganss et al. (2004b). This may be due to various reasons, including the fact that we have used a smoothed surface density distribution and are considering a larger opening angle than was done in Wambsganss et al. (2004b). However, our results seem to be in good agreement with those quoted in Dalal et al. (2005), who report a secondary plane contribution of $\lesssim 7\%$ for a source redshift, $z_s = 2.0$ and angular scales $\gtrsim 10''$.

5. ERROR IN THE MASS ESTIMATION OF THE MAIN LENS

As discussed in the previous section, there is a finite and non-negligible probability that more than one lens is responsible for causing a strong lensing event. We may ask, by how much we can underestimate the mass in lensing cluster or galaxy by assuming that it is the only object responsible for lensing. We study this by again shooting rays through the planes and considering cases where the total convergence along the line of sight, $\kappa_{\text{tot}} \equiv \sum \kappa_i$ is greater than unity. Within these cases, we look at the difference between κ_{tot} and the maximum value, κ_{\max} among the κ_i 's on the planes in front of the source. This maximum value would be the true value of the convergence due to the primary cluster which is the major contributor to the lensing. If we think that the other planes do not contribute significantly, we would assign the value κ_{tot} to it instead of κ_{\max} and this will be an error because statistically the contribution due to secondary planes is not negligible. To obtain an estimate of the error, we repeat the exercise several times and draw up the probability distribution of the percentage error in κ , defined as,

$$\epsilon = \frac{\kappa_{\text{tot}} - \kappa_{\max}}{\kappa_{\max}} \times 100. \quad (37)$$

We do this for a source redshifts of 2.0 and 3.0 and for each redshift, draw up two probability distributions of ϵ corresponding to the smoothing scales of $7.5''$ and $30''$. The results are displayed in Fig. 9. The distributions for the smaller smoothing length ($\sim 7.5''$) has a much longer tail (not shown), extending to very high values of ϵ , than that for the larger smoothing angle. This is expected because the probability density function of κ has a heavier tail for small smoothing angles, than it has for larger ones. We have also attempted to quantify the central value and the dispersion in ϵ by fitting a Lorentzian-like function to the data, given by,

TABLE 1
PARAMETERS OBTAINED BY FITTING THE FORMULA IN EQUATION (38) TO THE PDF OF ϵ .

Source redshift (z_s)	Peak centroid (a_1)		HWHM (a_2)	
	$\theta = 5''$	$\theta = 20''$	$\theta = 5''$	$\theta = 20''$
2.0	1.52	4.82	4.93	4.39
3.0	1.45	4.55	6.52	4.62

$$f(x) = \frac{a_0}{[(x - a_1)/a_2]^2 + 1} + a_3 x, \quad (38)$$

where the best fit values of the parameters a_1 and a_2 are interpreted as the centroid and the dispersion (half-width at half-maximum, or HWHM) of the distribution. These values are tabulated in Table 1.

From the table we note that the dispersions in the values of ϵ around the peak increases with increasing source redshift for a fixed smoothing angle. This is expected because for higher redshifts there are more intervening planes and hence a larger range of masses in clumps lining up with the main lens. Also, at a given redshift, the spread in the values of ϵ is larger for a smaller smoothing angle. This, basically reflects the fact that the variance of κ is higher for smaller smoothing angles. We also note, that although the distributions for the smaller smoothing angle has a much longer tail (not shown), the most probable error corresponding to it is lower than that for the larger smoothing angle. This basically means that although multiple-plane cases are less probable for the larger smoothing angle, when they happen, they carry a larger contribution of mass from the auxiliary planes, than in a typical small angle case. For instance, in a two-plane case the value of κ on the secondary plane typically come from the region around the peak of the PDF of κ on that plane. For a small smoothing angle, the peak is close to $\kappa \lesssim 0$ (the ray mostly hits slightly underdense regions), whereas as the smoothing angle grows the peak moves towards more positive values of κ . So in a supercritical line of sight, secondary planes will typically contribute higher values of κ when the smoothing angle is larger. This also applies to all the planes along the line of sight in general, so that for a larger smoothing angle the residual sum from all other planes, $\kappa_{\text{tot}} - \kappa_{\text{max}}$, will be higher than that for a smaller smoothing angle.

6. EFFECT OF INHOMOGENEITIES ON ANGULAR DIAMETER DISTANCES

Inhomogeneities in the distribution of matter induce fluctuations in the measured value of the angular diameter distance at a fixed redshift, z . The problem has been attacked directly from general relativity (Zel'Dovich 1964; Refsdal 1970; Dyer & Roeder 1974; Sasaki 1987; Fudamase & Sasaki 1989; Watanabe & Tomita 1990; Segal 1993; Kantowski et al. 1995; Linder 1998; Kantowski et al. 2000; Mörtsell 2002; Demianski et al. 2003) producing results of varied complexity.

In the following, we shall show that under certain simplifying assumptions, the differential equation for angular diameter distances can be obtained as a continuum limit to multiple gravitational lensing by successive

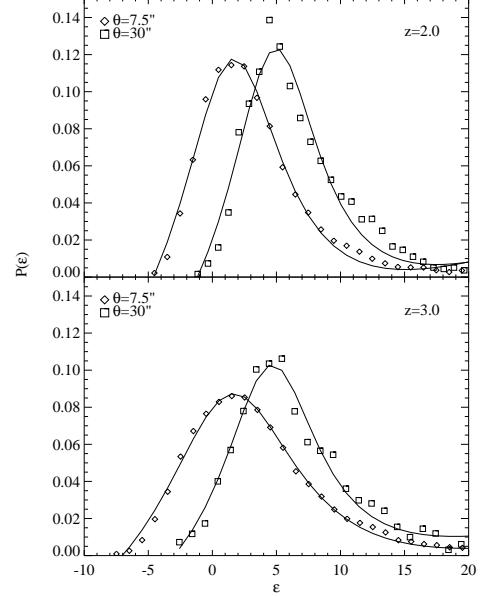


FIG. 9.— Probability distribution of the percentage error, ϵ , in the mass of the main lens for two different smoothing angles. *Top:* Source redshift, $z_s = 2.0$. *Bottom:* Source redshift, $z_s = 3.0$.

mass-sheets. This will provide a framework for studying angular-diameter distances in an inhomogeneous universe as a multiple deflection problem. The calculations are much in the spirit of Hadrović & Binney (1997), although we have generalized their results to include a non-zero cosmological constant.

Let us use the convention that \tilde{D} and D , respectively, denote angular-diameter distance before (empty beam) and after lensing is taken into account.

Fig. 10 shows a light ray that passes two mass sheets at impact parameters (measured with respect to the fiducial ray, OO') ξ_1 and ξ_2 and is deflected through angles α_1 and α_2 , respectively. A simple geometrical argument (see, for e.g. Hadrović & Binney 1997) shows that the impact parameter, ξ_j on the j -th plane is related to ξ_1 through the relation,

$$\xi_j = \frac{\tilde{D}_j}{\tilde{D}_1} \xi_1 - \sum_{i=1}^{j-1} \tilde{D}_{ij} \alpha_i(\xi_i), \quad (39)$$

where \tilde{D}_i is the empty-beam angular diameter distance to the i -th plane and \tilde{D}_{ij} denotes the empty-beam angular diameter distance between the i -th and the j -th planes.

If we assume that the mass distribution inside the beam on each mass sheet is uniform, with a surface density Σ , then the angle of deflection is given by,

$$\alpha = \frac{4\pi G}{c^2} \Sigma \xi. \quad (40)$$

Hence in this case the impact parameters are all parallel and satisfy

$$\xi_j = \frac{\tilde{D}_j}{\tilde{D}_1} \xi_1 - \frac{4\pi G}{c^2} \sum_{i=1}^{j-1} \Sigma_i \tilde{D}_{ij} \xi_i. \quad (41)$$

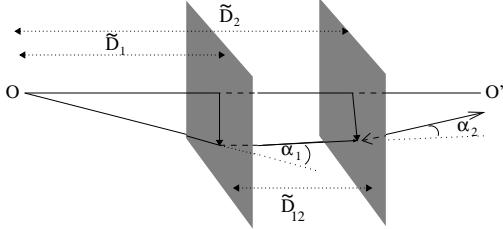


FIG. 10.— Lensing by a series of mass sheets.

The angular diameter distance to an object with diameter $\xi_i = |\xi_i|$ is $D_i \equiv \xi_i / (\xi_1 / \tilde{D}_1)$. Therefore, on taking the scalar form of equation (41) and dividing throughout by (ξ_1 / \tilde{D}_1) we find that the equation satisfied by the true angular diameter distances D_j has the form,

$$D_j = \tilde{D}_j - \frac{4\pi G}{c^2} \sum_{i=1}^{j-1} \Sigma_i D_i \tilde{D}_{ij}. \quad (42)$$

6.1. Equation for $\tilde{D}(y, z)$.

Let us denote the redshifts of the i -th and the j -th planes as y and z , respectively. Also, let the continuous version of the distance \tilde{D}_{ij} be denoted by $\tilde{D}(y, z)$.

$\tilde{D}(y, z)$ is the angular diameter distance in an empty beam (vanishing Ricci tensor and shear) and hence the focusing theorem (Schneider et al. 1992, equation (3.64)) takes the form,

$$\frac{d^2 \tilde{D}}{d\tau^2} = 0, \quad (43)$$

where τ is the affine parameter, defined through,

$$d\tau = \frac{dz}{(1+z)^2 \sqrt{\Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda}} \equiv \frac{dz}{g(z)}, \quad (44)$$

where we have defined,

$$g(z) \equiv (1+z)^2 \sqrt{\Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda}.$$

With the help of equation (44), equation (43) can be expressed in terms of redshift as,

$$g(z) \frac{\partial}{\partial z} \left(g(z) \frac{\partial}{\partial z} \tilde{D}(y, z) \right) = 0. \quad (45)$$

The initial condition is given by (Schneider et al. 1992, equation (4.53)), same arguments, but with $\Omega_\Lambda \neq 0$,

$$\frac{\partial \tilde{D}(y, z)}{\partial z} \Big|_{z=y} = \frac{(1+y)}{g(y)}. \quad (46)$$

6.2. Transition to continuum

Now we shall fill the beam with a cosmological matter density and use equation (42) to calculate the angular-diameter distance to a redshift z . We consider matter in the beam to be sliced up into discs of infinitesimal thickness. The disc that lies between redshifts $y + dy$ and y has surface density

$$d\Sigma = \rho_m(y) \frac{dr_{\text{prop}}}{dy} dy, \quad (47)$$

where, the proper distance, r_{prop} is given by,

$$\begin{aligned} dr_{\text{prop}} &= -cdt \\ &= \frac{c}{H_0} \frac{dy}{(1+y) \sqrt{\Omega_m(1+y)^3 + \Omega_k(1+y)^2 + \Omega_\Lambda}} \\ &= \frac{c}{H_0} \frac{(1+y)}{g(y)} \end{aligned} \quad (48)$$

and the physical density at redshift y , $\rho_m(y)$ is given by,

$$\rho_m(y) = \frac{3H_0^2}{8\pi G} \Omega_m [1 + \delta(y)] (1+y)^3, \quad (49)$$

where Ω_m is the dimensionless density parameter in the present universe and we have entertained the possibility of having inhomogeneous distribution of matter through the perturbation $\delta(y)$ in density.

Using equations (48), (49) and (47), we have,

$$\begin{aligned} d\Sigma &= \rho_m [1 + \delta(y)] (1+y)^3 \frac{c}{H_0} \frac{(1+y)}{g(y)} dy \\ &= \frac{3cH_0}{8\pi G} \left[\Omega_m \{1 + \delta(y)\} \frac{(1+y)^4}{g(y)} \right] dy. \end{aligned} \quad (50)$$

The continuum limit of the multiple deflection equation (42), after making the replacements: $\tilde{D}_j \rightarrow \tilde{D}(z)$, $D_i \rightarrow D(y)$, $\tilde{D}_{ij} \rightarrow \tilde{D}(y, z)$ and $\sum_{i=1}^{j-1} \rightarrow \int_0^z dy$, takes the form,

$$D(z) = \tilde{D}(z) - \frac{4\pi G}{c^2} \int dy \frac{d\Sigma}{dy} D(y) \tilde{D}(y, z). \quad (51)$$

Using equation (50) and expressing all distances in units of c/H_0 , the above equation becomes,

$$D(z) = \tilde{D}(z) - \frac{3}{2} \Omega_m \int_0^z dy \frac{[1 + \delta(y)] (1+y)^4}{g(y)} D(y) \tilde{D}(y, z) \quad (52)$$

We can convert this integral equation to a differential one as follows. Upon differentiating equation (52) with respect to z , we have,

$$\begin{aligned} \frac{dD}{dz} &= \frac{d\tilde{D}}{dz} \\ &\quad - \frac{3}{2} \Omega_m \int_0^z dy [1 + \delta(y)] \frac{(1+y)^4}{g(y)} D(y) \frac{\partial}{\partial z} \tilde{D}(y, z), \end{aligned}$$

and writing $\Omega_m[1 + \delta(y)]$ as $\Omega_m(y)$ for brevity, we further have,

$$\begin{aligned} \frac{d}{dz} \left(g \frac{dD}{dz} \right) &= \frac{d}{dz} \left(g(z) \frac{d\tilde{D}}{dz} \right) \\ &\quad - \frac{3}{2} \Omega_m(z) (1+z)^4 D(z) \left(\frac{\partial}{\partial z} \tilde{D}(y, z) \right)_{y=z} \\ &\quad - \frac{3}{2} \Omega_m \int_0^z dy \left[\{1 + \delta(y)\} \frac{(1+y)^4}{g(y)} D(y) \right. \\ &\quad \left. \times \frac{\partial}{\partial z} \left(g(z) \frac{\partial}{\partial z} \tilde{D}(y, z) \right) \right]. \end{aligned}$$

The first and the last terms on the right hand side of the above equation vanishes by virtue of (45). Using (46), we finally have the differential equations for $D(z)$ as,

$$g(z) \frac{d}{dz} \left(g(z) \frac{dD}{dz} \right) + \frac{3}{2} \Omega_m(1 + \delta(z)) (1+z)^5 D(z) = 0. \quad (53)$$

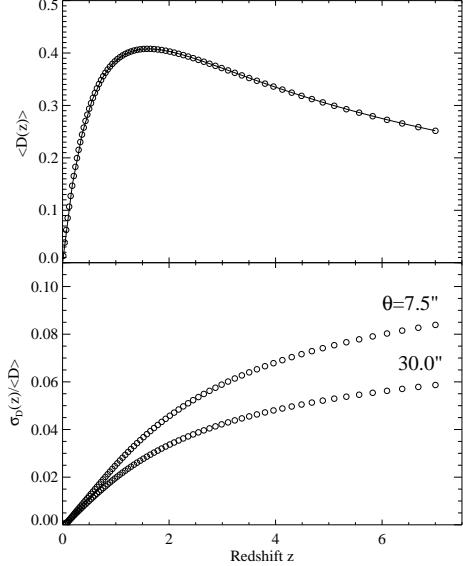


FIG. 11.— *Top*: The mean angular diameter distances obtained by Monte Carlo sampling of surface densities from the lens planes. The continuous line represents the smooth Friedman Robertson Walker (FRW) angular diameter distance as a function of redshift. The data are for an angular smoothing scale of $7.5''$. *Bottom*: The ratio, $\sigma_D(z)/\langle D(z) \rangle$, of the standard deviation to the mean of the angular diameter distance to each redshift, for two angular smoothing scales, $7.5''$ and $30.0''$.

If $\delta \equiv 0$ then equation (53) is the correct differential equation for angular diameter distance in an FRW universe. If the ray propagates through a smooth underdense region having a constant comoving density $\alpha\Omega_m$ ($0 < \alpha < 1$), then the equation reduces to the historical Dyer-Roeder equation.

Hence, we have shown that the angular diameter distances can be studied as multiple deflection lensing problem. It is important to discuss the simplifying assumptions that lead to this derivation. Firstly, we have assumed that the focusing of rays propagating over cosmological distances is unaffected by shear due to clumps of matter lying outside the beam. Futamase & Sasaki (1989) and Watanabe & Sasaki (1990) show that shear contributes negligibly to focusing as long as the scale of inhomogeneities are greater than or comparable to galactic scales. Since we study this problem with smoothing scales $\sim 10''$, this assumption seems to be a reasonable one. Secondly, we have assumed that matter that lies between the redshifts z and $z + dz$ forms a uniform disc. This is a nontrivial approximation but one that is particularly suited to our setup, in which we are looking at surface mass distributions on planes which have been already smoothed by a window of a certain angular size.

6.3. Implementation and results

For each beam, we approximate the continuous distribution with a Monte Carlo approach utilizing 38 values of Σ_θ from the PDFs corresponding to the 38 planes and compute the angular diameter distances out to various redshifts using equation (42). In the top panel Fig. 11, we show the result of a Monte Carlo run with 3×10^4

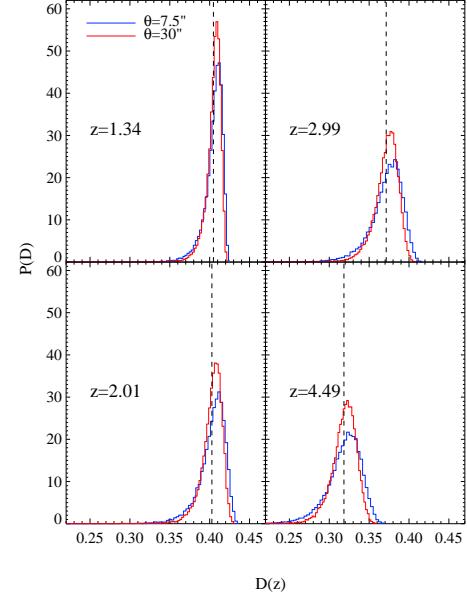


FIG. 12.— The probability distribution function of the angular diameter distance to four different redshifts, $z = 1.34, 2.01, 2.99$ and 4.49 and for two smoothing angles $\theta = 7.5''$ (blue histograms) and $30''$ (red histograms). The dashed vertical lines mark the position of the average or FRW angular diameter distance at each redshift.

beams, each having an opening angle of $7.5''$. As expected, the mean angular diameter distances perfectly match the smooth Friedman Robertson Walker (FRW) values. However, as displayed in the bottom panel of the same figure, the ratio of the standard deviation to the mean value of the angular diameter distance, increases with increasing redshift, amounting to a probable error of $\sim 5\%$ at a redshift of ~ 2.0 and rising up to $\sim 9\%$ for a redshift ~ 7.0 , for the smoothing angle of $7.5''$. The dispersions diminish with an increase in the smoothing angle, as is evident from the data points for the smoothing angle of $30''$. This is expected because with a larger smoothing angle, the surface densities encountered by the beam departs from the mean density with lesser probability. In Fig 12, we show the probability distributions of the angular diameter distance D , at four redshifts. These distributions are skewed towards values of D lower than the FRW values and the skewness is higher for the smaller smoothing angle. The skewness implies that a typical ray more frequently passes through an underdense region than it does an overdense one, and as discussed before in relation to the error in lens mass estimation, this bias becomes less important as the smoothing angle increases. Since various combinations of the angular diameter distance enter the calculations of lensing, the dispersion in its value, as displayed here, may systematically affect the use of lensing to determine of the properties of lenses or aspects of the underlying cosmology.

7. CONCLUSIONS

In the present paper, we have studied gravitational lensing with a multiple lens plane approach and have proposed a new analytic model for the dark matter convergence on the lens planes. Like the lognormal, this model

is entirely defined by $\langle \kappa^2 \rangle_\theta$, the variance of κ on a given lens plane and for a given smoothing angle θ . We showed that these variances can be theoretically predicted from the knowledge of the nonlinear power spectrum under a given cosmological model. Our tests indicate that the proposed model is far better than the lognormal in the ultra-nonlinear regime of small smoothing angles ($< 1'$) and describes the low-probability, high κ tail of the PDF of κ with quite high accuracy. We illustrated a number of possible applications of the model. With the PDF of κ on each plane accurately defined by our model, we could mimic an expensive ray shooting simulation to find out the importance of multiple clumps of matter in producing strong lensing, by a hugely faster Monte Carlo technique. We showed that the probabilities that a single plane on the line-of-sight is supercritical can be approximately obtained by a simple analytical calculation. We further demonstrated that secondary matter clumps are unimportant at source redshifts ~ 1 but may contribute up to about 15% of the single-plane cases at higher redshifts. The chance accumulation of significant amounts of secondary matter along the line-of-sight will therefore result in systematic inaccuracies in the determination of cluster masses, typically overestimating them by about 5% with in addition a $\sim 5 - 6\%$ standard deviation. We took a simplified model of the matter distribution inside the beam and were able to cast the problem of the determination of angular diameter distances into a multiple lensing problem. We could then use our analytic model to estimate the dispersions in angular diameter distances due to inhomogeneities in the matter field. As a cautionary note, we would like to add that the for lensing involving small angular scales, the role of baryons become important and the predictions from our model in this regime, should therefore be looked at as some kind of a lower limit. An important future improvement to our model would therefore be the inclusion of the effect of baryons. It should be noted that one important advantage of our method over the halo model approach is the ability to compute strong lensing by multiple lens planes, which is rather difficult in the halo approach. Therefore, our approach provides a complementary analytic tool to the halo model based calculations, in the sense that our model will be useful to examine the effect of secondary matter when combined with the halo model approach.

There are several ways in which the power of this model could be harnessed. As noted briefly at the end

of Section 3, this model may pave the way for a more accurate description of the line-of-sight-integrated convergence in the small angle regime where the lognormal approximation is sometimes inadequate. This issue is now in progress and will be presented elsewhere. It will also be of interest to study whether a similar model describes the 3-D PDF of dark matter density at very small scales. As the model is fixed entirely by the cosmology, one could test how efficiently it can be used to probe cosmological parameters. One of the several possible studies along these lines could be to determine how the predictions are altered by a different model of dark energy like the quintessence (Caldwell et al. 1998). With the non-Gaussianities in the lensing magnification well described by this model, one can use the scatter in Type IA supernovae to better constrain parameters like σ_8 (Dodelson & Vallinotto 2005; Cooray et al. 2005) or use multiple lens systems to better estimate the Hubble parameter H_0 (Fassnacht et al. 2005).

The statistics of gravitational lensing encodes the statistical features of matter distribution and hence probes cosmology at both linear and nonlinear scales. We expect that the model proposed in this paper will be able to play an important role in learning more about cosmology in a fashion complementary to that derivable from the analysis of large numerical simulations. The latter, of course, permit more detailed questions to be raised and answered than the analytical approach presented here. But the present approach is of sufficient accuracy and so much more computationally efficient than than based on direct simulation that we can now efficiently compare a suite of cosmological models. As an example of this, we could ask for the differences in the gravitational lensing predicted for models which give essentially identical, i.e. degenerate, cosmic microwave background signals (see, for e.g. Huey et al. 1999; Maor et al. 2001, 2002).

We sincerely thank Paul Bode for providing the data from the TPM simulations. SD specifically acknowledges his help in explaining various aspects of the simulation and the data. We thank Joe Hennawi for several careful readings of the manuscript and useful comments. SD would also like to thank Masamune Oguri for useful discussions and comments. Finally, we would like to thank the referee for his thoughtful suggestions.

REFERENCES

Augusto, P., et al. 2001, MNRAS, 326, 1007
 Barber, A. J., Munshi, D., & Valageas, P. 2004, MNRAS, 347, 667
 Barber, A. J., & Taylor, A. N. 2003, MNRAS, 344, 789
 Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
 Bartelmann, M., Huss, A., Colberg, J. M., Jenkins, A., & Pearce, F. R. 1998, A&A, 330, 1
 Bartelmann, M., & Schneider, P. 2001, Phys. Rep., 340, 291
 Belokurov, V. A., Shimanovskaya, E. V., Sazhin, M. V., Yagola, A. G., Artamonov, B. P., Shlyapin, V. N., & Khamitov, I. 2001, Astronomy Reports, 45, 759
 Bernardeau, F., van Waerbeke, L., & Mellier, Y. 1997, A&A, 322, 1
 Blandford, R. D., & Narayan, R. 1992, ARA&A, 30, 311
 Bode, P. & Ostriker, J. P. 2003, ApJS, 145, 1
 Broadhurst, T., Huang, X., Frye, B., & Ellis, R. 2000, ApJ, 534, L15
 Cen, R., Gott, J. R. I., Ostriker, J. P., & Turner, E. L. 1994, ApJ, 423, 1
 Caldwell, R. R., Dave, R., & Steinhardt, P. J. 1998, Physical Review Letters, 80, 1582
 Chae, K.-H., Mao, S., & Augusto, P. 2001, MNRAS, 326, 1015
 Claeskens, J.-F., Lee, D.-W., Remy, M., Sluse, D., & Surdej, J. 2000, A&A, 356, 840
 Claeskens, J. -, Sluse, D., Riaud, P., & Surdej, J. 2006, ArXiv Astrophysics e-prints, arXiv:astro-ph/0602309
 Cohn, J. D., Kochanek, C. S., McLeod, B. A., & Keeton, C. R. 2001, ApJ, 554, 1216
 Cooray, A., Huterer, D., & Holz, D. 2005, ArXiv Astrophysics e-prints, arXiv:astro-ph/0509581
 Dalal, N., Hennawi, J. F., & Bode, P. 2005, ApJ, 622, 99
 Dalal, N., Holder, G., & Hennawi, J. F. 2004, ApJ, 609, 50
 Demianski, M., de Ritis, R., Marino, A. A., & Piedipalumbo, E. 2003, A&A, 411, 33

Dodelson, S. 2003, Modern cosmology / Scott Dodelson. Amsterdam (Netherlands): Academic Press. ISBN 0-12-219141-2, 2003, XIII + 440 p.

Dodelson, S., & Vallinotto, A. 2005, ArXiv Astrophysics e-prints, arXiv:astro-ph/0511086

Dyer, C. C., & Roeder, R. C. 1974, ApJ, 189, 167

Fassnacht, C. D., Gal, R. R., Lubin, L. M., McKean, J. P., Squires, G. K., & Readhead, A. C. S. 2005, ArXiv Astrophysics e-prints, arXiv:astro-ph/0510728

Futamase, T., & Sasaki, M. 1989, Phys. Rev. D, 40, 2502

Hadrović, F., & Binney, J. 1997, ArXiv Astrophysics e-prints, arXiv:astro-ph/970811

Hennawi, J. F., Dalal, N., & Bode, P. 2005, ArXiv Astrophysics e-prints, arXiv:astro-ph/0506175

Hennawi, J. F., Dalal, N., Bode, P., & Ostriker, J. P. 2005, ArXiv Astrophysics e-prints, arXiv:astro-ph/0506171

Ho, S., & White, M. 2004, ArXiv Astrophysics e-prints, arXiv:astro-ph/0408245

Huey, G., Wang, L., Dave, R., Caldwell, R. R., & Steinhardt, P. J. 1999, Phys. Rev. D, 59, 063005

Inada, N., et al. 2003, Nature, 426, 810

Jain, B. 2002, ApJ, 580, L3

Jain, B., & Seljak, U. 1997, ApJ, 484, 560

Johnston, D. E., et al. 2003, AJ, 126, 2281

Kaiser, N. 1998, ApJ, 498, 26

Kaiser, N. 1992, LNP Vol. 408: New Insights into the Universe, 408, 247

Kantowski, R., Kao, J. K., & Thomas, R. C. 2000, ApJ, 545, 549

Kantowski, R., Vaughan, T., & Branch, D. 1995, ApJ, 447, 35

Kayo, I., Taruya, A., & Suto, Y. 2001, ApJ, 561, 22

Keeton, C. R. 2001, ApJ, 561, 46

Kormann, R., Schneider, P., & Bartelmann, M. 1994, A&A, 286, 357

Li, G. L., Mao, S., Jing, Y. P., Bartelmann, M., Kang, X., & Meneghetti, M. 2005, ArXiv Astrophysics e-prints, astro-ph/0503172

Li, L.-X., & Ostriker, J. P. 2002, ApJ, 566, 652

Li, L.-X., & Ostriker, J. P. 2003, ApJ, 595, 603

Linder, E. V. 1998, ApJ, 497, 28

Ma, C.-P., & Fry, J. N. 2000, ApJ, 531, L87

Maor, I., Brustein, R., McMahon, J., & Steinhardt, P. J. 2002, Phys. Rev. D, 65, 123003

Maor, I., Brustein, R., & Steinhardt, P. J. 2001, Physical Review Letters, 87, 049901

Mörtsell, E. 2002, A&A, 382, 787

Nakamura, T. T., & Suto, Y. 1997, Progress of Theoretical Physics, 97, 49

Narayan, R., & Blandford, R. 1991, New York Academy Sciences Annals, 647, 117

Oguri, M., & Keeton, C. R. 2004, ApJ, 610, 663

Oguri, M., Lee, J., & Suto, Y. 2003, ApJ, 599, 7

Oguri, M., Taruya, A., Suto, Y., & Turner, E. L. 2002, ApJ, 568, 488

Patnaik, A. R., & Narasimha, D. 2001, MNRAS, 326, 1403

Peacock, J. A., & Dodds, S. J. 1996, MNRAS, 280, L19

Petters, A. O., Levine, H., & Wambsganss, J. 2001, Singularity theory and gravitational lensing / Arlie O. Petters, Harold Levine, Joachim Wambsganss. Boston : Birkhäuser, c2001. (Progress in mathematical physics ; v. 21),

Premadi, P., Martel, H., Matzner, R., & Futamase, T. 2001, ApJS, 135, 7

Premadi, P., & Martel, H. 2004, ApJ, 611, 1

Refregier, A. 2003, ARA&A, 41, 645

Refsdal, S. 1970, ApJ, 159, 357

Sasaki, M. 1987, MNRAS, 228, 653

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses, XIV, 560 pp. 112 figs.. Springer-Verlag Berlin Heidelberg New York.

Schneider, P., van Waerbeke, L., Jain, B., & Kruse, G. 1998, MNRAS, 296, 873

Segal, I. E. 1993, Proceedings of the National Academy of Science, 90, 4798

Seljak, U. 2000, MNRAS, 318, 203

Shapiro, P. R., & Iliev, I. T. 2000, ApJ, 542, L1

Sluse, D., et al. 2005, ArXiv Astrophysics e-prints, arXiv:astro-ph/0512473

Smith, R. E. et al. 2003, MNRAS, 341, 1311

Spergel, D. N., et al. 2003, ApJS, 148, 175

Stebbins, A. 2004, ArXiv Astrophysics e-prints, arXiv:astro-ph/9609149

Sugiyama, N. 1995, ApJS, 100, 281

Taruya, A., Takada, M., Hamana, T., Kayo, I., & Futamase, T. 2002, ApJ, 571, 638

Tonry, J. L. 1998, AJ, 115, 1

Turner, E. L., Ostriker, J. P., & Gott, J. R. 1984, ApJ, 284, 1

van Waerbeke, L., Bernardeau, F., & Mellier, Y. 1999, A&A, 342, 15

Villumsen, J. V. 1996, MNRAS, 281, 369

Wambsganss, J., Bode, P., & Ostriker, J. P. 2004, ApJ, 606, L93

Wambsganss, J., Bode, P., & Ostriker, J. P. 2004, ArXiv Astrophysics e-prints, astro-ph/0405147

Watanabe, K., & Sasaki, M. 1990, PASJ, 42, L33

Watanabe, K., & Tomita, K. 1990, ApJ, 355, 1

Wayth, R. B., Warren, S. J., Lewis, G. F., & Hewett, P. C. 2005, MNRAS, 360, 1333

Zel'Dovich, Y. B. 1964, Soviet Astronomy, 8, 13